

# PROJECT FINAL REPORT

**Grant Agreement number: 288956**

**Project acronym: NADINE**

**Project title: New tools and Algorithms for Directed Network analysis**

**Funding Scheme: Small or medium-scale focused research project (STREP)**

**Period covered: from 01.05.2012 to 30.04.2015**

**Name of the scientific representative of the project's co-ordinator<sup>1</sup>, Title and Organisation:**

**Dr. Dima Shepelyansky**

**Directeur de recherche au CNRS**

**Lab de Phys. Theorique, Universite Paul Sabatier, 31062 Toulouse, France**

**Tel: +331 5 61556068, Fax: +33 5 61556065, Secr.: +33 5 61557572**

**E-mail: [dima@irsamc.ups-tlse.fr](mailto:dima@irsamc.ups-tlse.fr)**

**Project website address: [www.quantware.ups-tlse.fr/FETNADINE](http://www.quantware.ups-tlse.fr/FETNADINE)**

---

1

## 4.1 Final publishable summary report

### An executive summary

On the scale of the past ten years, modern societies have developed enormous communication and social networks. Their classification and information retrieval becomes a formidable task for the society. Various search engines have been developed by private companies which are actively used by Internet users. Due to the recent enormous development of World Wide Web and communication networks, new tools and algorithms should be invented to characterize the properties of these networks on a more detailed and precise level. It is also highly important to have new tools to classify and rank enormous amount of network information in a way adapted to internal network structures and characteristics. The project will develop new algorithms to facilitate classification and information retrieval from large directed networks, including PageRank and CheiRank with two-dimensional ranking proposed by partners, using newly developed Monte Carlo methods. The Google matrix formed by the links of the network will be analyzed by analytical tools of Stochastic Processes, Random Matrix Theory and quantum chaos and by efficient numerical methods for large matrix diagonalization including the Arnoldi method. New tools and algorithms produced by the project will create fundamental basis for developers of new types of search and social media services, which will put Europe on leading positions in this important area dominated at present by other countries. NADINE tools will find applications in modern networks, including mobile communication networks which will play more and more important role in future. New characterization of complex networks will allow stakeholders to manage information extraction for social networks, communication and other networks in an efficient and rapid way. The project will create efficient voting systems in social networks that will pave the way for new types of democracy solutions in societies at a high communication level.

The central aims of this project are to develop new algorithms to facilitate classification and information retrieval from large directed networks, including PageRank and CheiRank with two-dimensional ranking proposed by partners, using newly developed Monte Carlo methods. The Google matrix formed by the links of the network is analyzed by analytical tools of Stochastic Processes, Random Matrix Theory and quantum chaos and by efficient numerical methods for large matrix diagonalization including the Arnoldi method. The investigations of real directed networks performed by the project highlight their new characteristics allowing to understand in a deeper way the hidden features of these networks. New tools and algorithms produced by the project create fundamental basis for developers of new types of search and social media services. The consortium has interdisciplinary skills since it unites partners from different sciences including physics, mathematics and computer science.

### A summary description:

WP1: The main objective of this work package is to lay mathematical foundations for development and application of new ranking schemes such as 2DRanking, and provide fast algorithms for their computation. PageRank is widely applied for ranking of nodes in directed networks including World Wide Web and citation graph. However, up to date, very little is known about mathematical properties of the resulting PageRank vector. The results of the consortium prove that the power law behaviour of PageRank is defined by the distribution of the in-degree. However, the dependence between these two quantities is remarkably different, e.g., for Web and Wikipedia. The partners also found that correlations of PageRank and CheiRank are small in some networks (e.g. for Linux kernel software and gene networks), and large in others (e.g. Web samples and Wikipedia). We will use novel methods, proposed by the consortium, to adequately measure correlations between node parameters, and obtain analytical description for 2DRanking, where these correlations are taken into account. We will extend our analysis to new centrality measures, of which desirable properties for specific network structures and applications will be justified by a mathematical model. Finally, our objective is to develop efficient Monte Carlo algorithms for evaluating centrality measures. Our results prove that such methods are remarkably efficient if the goal is to evaluate the ranking order, and not the exact values of

centrality scores. Our aim is to evaluate the required computational complexity of Monte Carlo in order to produce an informative ranking order.

WP2: WP2 investigates spectrum of Google matrices of such real networks as WWW university networks, network of hyperlinks between Wikipedia English articles, network of links of procedure call procedures in open source software. The Arnoldi method applied to the Linux network established the validity of fractal Weyl law, found recently in systems of quantum chaotic scattering and Perron-Frobenius operators of dynamical systems. WP2 investigates the spectrum of Wikipedia network analyzed recently. The eigenmodes, with eigenvalue modulus being close to the damping factor, correspond to slow relaxation modes in networks. Such modes should be linked with specific communities hidden inside network. The Arnoldi method allows to detect such modes in an effective way thus open new possibilities for extracting of hidden communities from networks. Examples of Google matrix spectrum for the WWW of Cambridge and Oxford Universities, obtained by the Arnoldi method, show decomposition of degenerate subspaces at damping factor equal unity. The size of degenerate subspaces can be rather large (around 40000 for the case of WWW of Cambridge University of total size 200000). This Arnoldi approach will be also applied for networks of Wikipedia articles, open source software networks, university networks. Fractal dimensions of the networks will be also determined. The Arnoldi method will be also used to detect communities, linked to eigenstates with eigenvalue close to one, in the Wikipedia articles network of  $N=3282257$  nodes extending previous results. The high efficiency of the Arnoldi method allows to handle Google matrices of very large size using modern computers available to the consortium. Delocalization properties of eigenstates will also be determined for various networks.

WP3: *Voting* is a basic decision procedure by which individuals express their preferences among a set of choices. Given the preferences of all voters (each one a permutation of the possible choices), a voting system generates a single choice. Voting theory studies how to select such a choice under certain optimization constraints. In particular, choices can be individuals that must be chosen for some purpose (e.g., to take a decision, or to represent the population). In *direct democracy*, each individual can vote any other individual. Recently, to obviate the lack of acquaintance between voter and voted individual in large populations, *liquid democracy* (a.k.a. *proxy voting*) has been introduced. In this case, a vote is given to some other individual that can keep it (and then we can perform an election just by majority) or give it away to someone else.

In social networks representing acquaintances between people (e.g., Facebook), however, we have a much more interesting scenario, as we are given from the start, for each individual, a set of users that are directly known (its neighbours in the graph). By restricting the ability to vote to acquaintances, we can obviate (even for very large networks) the problem of low representativity: if we give our vote to one of our acquaintances, we judge it apt to take a decision for us. Due to the large size of social networks (Facebook has currently more than 700 million active users), however, a direct application of liquid democracy can lead to a number of problems, most notably the loss of control of our vote: due to the small-world phenomenon, in a very small number of passages our vote can reach essentially any individual.

Recently, *viscous democracy* has been proposed for social networks by members of the consortium. Voters can only choose one of their neighbours, generating a *voting graph*—a directed graph of constant outdegree one. Each vote is passed to the chosen neighbour, but weakened by a multiplicative attenuation factor. If the vote travels too far, it is ineffective. It turns out that this is equivalent to computing Katz's index (or, in this case, due to the fixed outdegree of the graph, PageRank) on the voting graph—hence the name *spectral voting* for this kind of technique. Due to the known connection between path-based ranking and eigenvector-based ranking, the resulting scores turn out to be given by the dominant eigenvector of a suitable matrix

WP4: Methods of WP1-WP3 are implemented in large scale applications based on real data collected in WP5. Achievements in this WP are measured in terms of: the size of the data processed, with WP targets at Web scale, billions of objects; another benchmark is the approximation error of the fingerprinting and lazy update procedures, with the target to keep the error below the limit of notice in a user application. Special distributed network technologies will be developed to reach such goals. Spam filtering protocols will also be developed and tested. Using these tools and those of WP1-WP3, statistical analysis will be done for several types of important networks including Wikipedia in English, French, German, Italian and Spanish at different moments

of time evolution; open software procedure networks, genes and other networks. Applications of centrality measures to game theory will also be developed. We will generalize recent results for the Google matrix of world trade network to the case of multiproduct trade for which the matrix size is increased by two or more orders of magnitude.

WP5: WP5 develops efficient protocols for large scale network analysis and generates database collections that will be treated by the methods developed in WP1-WP3. To this aim specific skilful crawlers will be developed to collect information from modern enormous data bases. Data sets evolving in time will be analyzed by specially developed protocols.

## A description of the main S&T results/foregrounds:

**Highlights in the first reporting period** include work re on spectral properties of spectrum and eigenstates of Wikipedia and their links with communities detection [1], multilingual ranking of world persons with global heroes like Napoleon and Michael Jackson [2], design of new correlation measures between degrees of neighboring nodes [3], detailed analysis of Last.fm network of users [4], demonstration of four degrees of separation for the whole Facebook network of users [5].

[1] L.Ermann, K.M.Frahm and D.L.Shepelyansky, "Spectral properties of Google matrix of Wikipedia and other networks", Eur. Phys. J. B v.86, p.193 (2013), arXiv:1212.1068 [cs.IR]

[2] Y.-H.Eom and D.L.Shepelyansky, "Highlighting entanglement of cultures via ranking of multilingual Wikipedia articles", to appear in PLoS ONE Oct 2013, arXiv:1306.6259 [cs.SI]

[3] N. Litvak, and R. van der Hofstad "Uncovering disassortativity in large scale-free networks", Phys. Rev. E v.87, p.022801 (2013) (arXiv:1204.0266[physics.soc-ph])

[4] R. Pálovics, A.A. Benczúr. Temporal influence over the Last.fm social network. The 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining ASONAM 2013 Niagara Falls, Canada, August 25-28, 2013

[5] SebaL.Backstrom, P.Boldi, M.Rosa, J.Ugander, and S.Vigna. "Four degrees of separation", ACM Web Science 2012: Conference Proceedings, June 2012, pages 45-54, ACM Press (2012); best paper award, highlighted by the New York Times; (arXiv:1111.4570, 2012)

**Highlights in the second reporting period** include work on interactions of cultures and top 100 people of Wikipedia from ranking of 24 language editions [69], Google matrix analysis of the multiproduct world trade network [39], development of Monte Carlo algorithm for quick detection of high-degree entities in large directed networks [50], results for RecSys Challenge 2014: an ensemble of binary classifiers and matrix factorization [53], construction of a weighted correlation index for rankings with ties [70] (Refs numbers are given for the whole grant period as on <http://www.quantware.ups-tlse.fr/FETNADINE/pub.html> )

1.1H. [69] P4.16 Young Ho Eom, Pablo Aragon, David Laniado, Andreas Kaltenbrunner, Sebastiano Vigna, and Dima L. Shepelyansky, "**Interactions of cultures and top people of Wikipedia from ranking of 24 language editions**", PLoS ONE v.10(3), p.e0114825 (2015) (arXiv:1405.7183[cs.SI], 2014)

2.1H. [39] P1.19 L.Ermann and D.L.Shepelyansky, "**Google matrix analysis of the multiproduct world trade network**", Eur.Phys. J. B v.88, p.84 (2015) (arXiv:1502.00584[cond-mat.dis-nn])

3.2H. [50] P2.13 K.Avrachenkov, N.Litvak, L.Ostroumova-Prokhorenkova and E.Suyargulova, "**Quick detection of high-degree entities in large directed networks**", IEEE International Conference on Data Mining (ICDM 2014), 14-17 Dec 2014, Shenzhen, China. pp. 20-29. IEEE Computer Society (2014) (arXiv:1410.0571v2[cs.SI])

4.2H. [53] P3.9 R.Palovics, F. Ayala-Gomez, B. Csikota, B.Daroczy, L. Kocsis, D. Spadacene, A.A. Benczur, "**RecSys Challenge 2014: an ensemble of binary classifiers and matrix factorization**", Proceedings of the 2014 Recommender Systems Challenge (p. 13) ACM (2014)

5.2H. [70] P4.17 Sebastiano Vigna, "**A weighted correlation index for rankings with ties**", Proceedings of the 24th international conference on World Wide Web, ACM (2015) (arXiv:1404.3325[cs.SI], 2014)

Joint publications for the whole grant period are: [8] P1.8 (P1 and P3); [49] P2.12 (P2 and P4); [60] P3.16 (P1 and P3); [61] P3.17 (P2 and P3); [69] P4.16 (P1 and P4).

The project has fulfilled all deliverables and milestones for the reporting period and has resulted in collaborations between all partners. In total, since the beginning of the project, 73 papers and preprints have appeared during the period within the framework of NADINE. These works include 3 papers in PLoS ONE and 1 in Sci. Reports. The PLOS paper in 2015 of P1+P4 has been highlighted by Guardian, Independent, Le Figaro, EC CORIDS and press of about 20 countries. One paper of P4 had been highlighted by New York times. The Sci. Reports paper has been highlighted by Independent and MIT Tech. Rev. During the NADINE period the results have been reported on 67 international conferences.

## The potential impact

The results obtained during the grant period are presented in publications [1-73] available for public at the web page

<http://www.quantware.ups-tlse.fr/FETNADINE/pub.html>

Among press releases we note those for publication P4.16 highlighted by The Guardian, The Washington Post, Uppsala Universitet, EC Cordis, CNRS INP news (this CNRS highlight got the largest number of views on the day of its publication which exceeded the number of views for the CNRS highlight about Nobel prize for S.Haroche), in total there are about 20 countries highlighted this publication (see all links at the above web page).

The software developed during the grant period is publicly available at the web page

<http://www.quantware.ups-tlse.fr/FETNADINE/software.htm>

The algorithms and software developed by NADINE attracted interests of various institutions and companies. Thus the NADINE algorithms for the multiproduct world trade are reported to United Nations COMTRADE New York; this database continues to provide to P1 a free access to this database. These algorithms have been also applied to the data collected by OECD-WTO in collaboration with the Chief statistician H.Escaith (WTO Geneve with a joint publication P1.22). These results based on OECD-WTO database for the world network of economic activities are presented to the International conference at Bank for International Settlements, Basel in Oct 2015 <http://www.bis.org/events/gfi.htm> . It is clear that the algorithms developed by NADINE for multiproduct world trade can be directly applied to the treatment of stability of financial flows between bank units. This world network of bank units is estimated to be of a rather small size (7000 nodes for Federal

Reserve of USA, 2000 nodes for all banks of Germany and about 50000 nodes for the whole world). In various talks NADINE Coordinator argued that the NADINE methods will be very useful for financial networks allowing to prevent the financial crises (invited talk at Deutsche Bundesbank, Zentrale, Makroprudenzielle Analysen, Frankfurt am main, 6 Feb 2012 (slides are at <http://www.quantware.ups-tlse.fr/dima/dlstalks/dlstalk26.pdf> ) and at presentation talk of

NADINE project at International Workshop on Search Computing, Brussels 25-26 Sept 2012

(video of talk is at [http://videlectures.net/searchcomputing2012\\_shepelyansky\\_nadine/](http://videlectures.net/searchcomputing2012_shepelyansky_nadine/) ).

However, banks prefer to hide their financial flows (even flows between unknown anonymized bank units are not available). NADINE Coordinator expressed the firm opinion that EC political wish can make a political pressure on banks allowing to get for scientists an information about financial flows between bank units (for example, for old periods of time, or/and for no-named bank units). This would allow to create the World Bank Web (WBW) and to protect the human society from bank crises after which countries and common people are obliged to compensate bank losses. At present the money transfers with SWIFT clearly show that the WBW really exists. The computer science and mathematical methods developed by NADINE would allow to supervise the financial flows on WBW and bring to the society the stability of world financial system. There is an appealing analogy with the WWW: its appearance in 1991 changed the world on a scale of 20 years, we argue that the creation of the WBW (with restricted access to authorized community of politicians and researchers) will bring the financial stability to the world. This direction can be developed in the frame of FAPLIDIN project submitted to the FET Open in March 2015.

Algorithm developed by P2 (publication P2.13) attracted interest of Yandex (Russia) which wanted to make a patent on the basis of this algorithm (two co-authors of this paper work at Yandex) but the paper had been presented at a conference at early stage and a patent could not be realized after that.

Algorithms developed by P3 attracted interest of NOMAO.com (paper P3.16 with P1), other results of P3 had been solicited by Ericsson Hungary, P2 had been invited to a workshop organized by Airbus in Toulouse in Dec 2014 to present their algorithms.

P4 developed various software tools used by Facebook (paper P4.1 highlighted by New York Times), P4 got also a large Google grant on software development.

The datasets collected during the grant period are available for public at

<http://www.quantware.ups-tlse.fr/FETNADINE/datasets.htm>

The results present at the FET NADINE meetings, conferences are available at

<http://www.quantware.ups-tlse.fr/complexnetworks2012/>

<http://www.quantware.ups-tlse.fr/FETNADINE/dnd2013/>

<http://www.quantware.ups-tlse.fr/FETNADINE/dnd2014/>

Materials of Summer schools organized by NADINE (including video lectures for students) are available at

<http://www.quantware.ups-tlse.fr/ecoleluchon2014/>

<http://www.quantware.ups-tlse.fr/ecoleluchon2015/> (in preparation)

The obtained analytical results and software give clear recipes for advanced analysis of complex directed networks. These tools are advertised to PME companies like NOMAO, Gravity etc. The open public crawler developed by NADINE is used by various research groups and companies. NADINE also developed contacts with the World Trade Organisation (Geneve) and UN COMTRADE (New York) analyzing their international trade and economic activities data by NADINE tools.

Future developments and applications of the results of NADINE are planned to be extended in the proposed FET Open project FAPLIDIN (Fundamentals and APPLications of DIrected Networks) submitted to FET in March 2015 (Proposal number: SEP-210260860)

Furthermore, project logo, diagrams or photographs illustrating and promoting the work of the project are available at the project website. All contacts are given there.

## 4.2 Use and dissemination of foreground

See for more details Section above “The potential impact”. Below we give links to publications, conferences, software, data sets etc.

### Section A (public)

- Template A1: List of all scientific (peer reviewed) publications relating to the foreground of the project. All publications of the project are publicly available at the project website <http://www.quantware.ups-tlse.fr/FETNADINE/pub.html>
- Template A2: List of all dissemination activities (publications, conferences, workshops, web sites/applications, press releases, flyers, articles published in the popular press, videos, media briefings, presentations, exhibitions, thesis, interviews, films, TV clips, posters). All conferences presentations are described at project website <http://www.quantware.ups-tlse.fr/FETNADINE/conf.html>; video lectures for students are available at <http://www.quantware.ups-tlse.fr/ecoleluchon2014/program.html>

Software developed during the project is described at <http://www.quantware.ups-tlse.fr/FETNADINE/software.htm>

Data sets collected during the project are available at <http://www.quantware.ups-tlse.fr/FETNADINE/datasets.htm>

## Section A (public)

This section includes two templates

- Template A1: List of all scientific (peer reviewed) publications relating to the foreground of the project.
- Template A2: List of all dissemination activities (publications, conferences, workshops, web sites/applications, press releases, flyers, articles published in the popular press, videos, media briefings, presentations, exhibitions, thesis, interviews, films, TV clips, posters).

These tables are cumulative, which means that they should always show all publications and activities from the beginning until after the end of the project. Updates are possible at any time.

TEMPLATE A1: LIST OF SCIENTIFIC (PEER REVIEWED) PUBLICATIONS, STARTING WITH THE MOST IMPORTANT ONES										
NO.	Title	Main author	Title of the periodical or the series	Number, date or frequency	Publisher	Place of publication	Year of publication	Relevant pages	Permanent identifiers <sup>1</sup> (if available)	Is/Will open access <sup>2</sup> provided to this publication?
1	<i>Spectral properties of Google matrix of Wikipedia and other networks</i>	Shepelya nsky	<i>Eur. Phys. J. B</i>	v.86, p.193 (2013)	EPS	EU	2013	v.86, p.193		Available on project website
2	<i>Highlighting entanglement of cultures via ranking of multilingual Wikipedia articles</i>	Shepelya nsky	PLoS One	v.8(10), p.e74554 (2013)	PloS ONE	WORLD	2013	v.8(10), p.e74554		Open access
3	<i>Uncovering disassortativity in large scale-free networks</i>	Litvak	Phys Rev E	v.87, p.022801 (2013)	APS	USA	2013	v.87, p.022801		Available on project website

<sup>1</sup> A permanent identifier should be a persistent link to the published version full text if open access or abstract if article is pay per view) or to the final manuscript accepted for publication (link to article in repository).

<sup>2</sup> Open Access is defined as free of charge access for anyone via Internet. Please answer "yes" if the open access to the publication is already established and also if the embargo period for open access is not yet over but you intend to establish open access afterwards.



4	Temporal influence over the Last.fm social network	Benczur		The 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining ASONAM 2013 Niagara Falls, Canada, August 25-28, 2013		Canada	2013			Available on project website
5	Four degrees of separation	Vigna		ACM Web Science 2012: Conference Proceedings, June 2012, pages 45-54, ACM Press (2012);		USA	2012			Available on project website
6	Interactions of cultures and top people of Wikipedia from ranking of 24 language editions	Shepelyansky, Vigna	PLoS ONE	v.10(3), p.e0114825 (2015)	Plos ONE	world	2015	v.10(3), p.e0114825		Open access
7	Google matrix analysis of the multiproduct world trade network	Shepelyansky	Eur. Phys. J. B	v.88, p.84 (2015)	EPS	EU	2015	v.88, p.84		Available on project website
8	Quick detection of high-degree entities in large directed networks	Litvak	IEEE International Conference	IEEE International Conference on Data		USA	2014			Available on project website

			nce on Data Mining (ICDM 2014), 14-17 Dec 2014, Shenzh en, China. pp. 20- 29. IEEE Comput er Society (2014)	Mining (ICDM 2014), 14-17 Dec 2014, Shenzhen, China. pp. 20- 29. IEEE Computer Society (2014)						
<b>9</b>	RecSys Challenge 2014: an ensemble of binary classifiers and matrix factorization	Benczu r	Proc. 2014 Recom mender Sys. Challen geACM (2014)	Proceedings of the 2014 Recommender Systems Challenge (p. 13) ACM (2014)	ACM	USA	2014			Available on project website
<b>10</b>	A weighted correlation index for rankings with ties	Vigna	Proc. 24th Int.Conf WWW ACM (2015)	Proceedings of the 24th international conference on World Wide Web, ACM (2015)	ACM	USA	2015			Available on project website
<b>11</b>	<b>Towards two-</b>	Shepely	J.Phys.	v.45, p.275101	IOP	UK	2012	v.45,		Available on

	<b>dimensional search engines</b>	ansky	A	(2012)				p.275101 (2012)		project website
12	<b>Ecological analysis of world trade</b>	Shepely ansky	Phys. Lett. A	v.377, p.250 (2013)	Elsevier	NL	2013	v.377, p.250 (2013)		Available on project website
13	<b>PageRank model of opinion formation on social networks</b>	Shepely ansky	Physica A	v.391, p.5779 (2012)	Elsevier	NL	2012	v.391, p.5779 (2012)		Available on project website
14	<b>PageRank of integers</b>	Shepely ansky	J.Phys. A	v.45, p.405101 (2012)	IOP	UK	2012	v.45, p.405101 (2012)		Available on project website
15	<b>Google matrix of Twitter</b>	Shepely ansky	Eur.Phys. J.B	v.85, p.355 (2012)	EPS	EU	2012	v.85, p.355 (2012)		Available on project website
16	<b>Spectral properties of Google matrix of Wikipedia and other networks</b>	Shepely ansky	Eur.Phys. J.B	v.86, p.193 (2013)	EPS	EU	2013	v.86, p.193 (2013)		Available on project website
17	<b>Google matrix analysis of DNA sequences</b>	Shepely ansky	PloS ONE	v.8(5), p. e61519 (2013)	PloS ONE	world	2013	v.8(5), p. e61519 (2013)		Open access
18	<b>Time evolution of Wikipedia network ranking</b>	Shepely ansky	Eur.Phys. J.B	v.86, p.492 (2013)	EPS	EU	2013	v.86, p.492 (2013)		Available on project website
19	<b>PageRank model of opinion formation on Ulam networks</b>	Shepely ansky	Phys. Lett. A	v.377, p.3119 (2013)	Elsevier	NL	2012	v.377, p.3119 (2013)		Available on project website
20	<b>Poincare recurrences and Ulam method for the Chirikov standard map</b>	Shepely ansky	Eur.Phys. J.B	v.86, p.322(2013)	EPS	EU	2013	v.86, p.322(2013)		Available on project website
21	<b>Google matrix of the citation network of Physical Review</b>	Shepely ansky	Phys. Rev. E	v.89, p.052814 (2014)	APS	USA	2014	v.89, p.052814 (2014)		Available on project website
22	<b>Google matrix analysis</b>	Shepely	Phys.	v.378, p.1932	Elsevier	NL	2014	v.378,		Available on

	<b>of C.elegans neural network</b>	ansky	Lett. A	(2014)				p.1932 (2014)		project website
23	<b>Poisson statistics of PageRank probabilities of Twitter and Wikipedia networks</b>	Shepely ansky	Eur.Phys. J.B	v.87, p. 93 (2014)	EPS	EU	2014	v.87, p. 93 (2014)		Available on project website
24	<b>Google matrix analysis of directed networks</b>	Shepely ansky	Rev.Mod. Phys.	accepted	APS	USA	2015	accepted		Available on project website
25	<b>Generalized friendship paradox in complex networks: the case of scientific collaboration</b>	Eom	Sci.Reports.	v.4, p.4603 (2014)	Nature	UK	2014	v.4, p.4603 (2014)		Open access
26	<b>Generalized friendship paradox in networks with tunable degree-attribute correlation</b>	Eom	Phys.Rev. E	v.90, p.022809 (2014)	APS	USA	2014	v.90, p.022809 (2014)		Available on project website
27	<b>More ordering and communities in complex networks describing the game of go</b>	Georgiot	Eur.Phys. J.B	v.87, p.246 (2014)	EPS	EU	2014	v.87, p.246 (2014)		ibid
28	<b>Google matrix analysis of the multiproduct world trade network</b>	Shepely ansky	Eur.Phys. J.B	v.88, p.84 (2015)	EPS	EU	2015	v.88, p.84 (2015)		ibid
29	<b>Anderson transition for Google matrix eigenstates</b>	Shepely ansky	Ann. der Physik (Berlin)	accepted	Springer	Germany	2015	accepted	DOI 10.1002/andp.201500110	ibid
30	<b>Opinion formation driven by PageRank node influence on directed networks</b>	Shepely ansky	Physica A	accepted	Elsevier	NL	2015	accepted		ibid
31	<b>Google matrix of the world network of economic activities</b>	Shepely ansky	Eur.Phys. J.B	accepted	EPS	EU	2015	accepted		ibid

32	<b>Degree-degree correlations in random graphs with heavy-tailed degrees</b>	Litvak	Internet Math.	v.10(3-4), p.287 (2014)	Taylor-Francis	EU	2014	v.10(3-4), p.287 (2014)		ibid
33	<b>Alpha current flow betweenness centrality</b>	Litvak	Lecture Notes in Computer Science	v.8305 p.106 (2013)	Springer	Germany	2013	v.8305 p.106 (2013)		ibid
34	<b>Degree-degree dependencies in directed networks with heavy-tailed degrees</b>	Litvak	Internet Math	v.11(2), p.155 (2015)	Taylor-Francis	EU	2015	v.11(2), p.155 (2015)		ibid
35	<b>Modelling of trends in Twitter using retweet graph dynamics</b>	Litvak	Lecture Notes in Computer Science	2014 (8882), Springer (2014)	Springer	Germany	2014	p.120 (8882), Springer (2014)		ibid
36	<b>Introduction to Special Issue on Searching and Mining the Web and Social Networks</b>	Litvak Vigna	Internet Math	v.10(3-4), p.219-221 (2014)	Taylor-Francis	EU	2014	v.10(3-4), p.219-221 (2014)		ibid
37	<b>The classification power of Web features</b>	Benczur	Internet Math	v.10(1), p.31 (2014)	Taylor-Francis	EU	2014	v.10(1), p.31 (2014)		ibid
38	<b>Temporal Wikipedia search by editr and linkage</b>	Benczur	SIGIR Conf			Ireland	2013			ibid
39	<b>Exploiting temporal influence in online recommendation</b>	Benczur	Proc. 8th ACM Conf on Recommender Sys.		ACM	USA	2015	p.273		ibid

40	<b>Statistical analysis of NOMAO customer votes for spots of France</b>	Benczur , Shepely ansky	Eur. Phys. J.B	accepted	EPS	EU	2015	accepted	ibid
41	<b>Four degrees of separation, really</b>	Vigna	IEEE/A CM Int. Conf. Adv.s Social Netw. Anall. Mining	IEEE, 2012: 1222-1227	IEEE	USA	2012	IEEE, 2012: 1222-1227	ibid
42	<b>Arc-Community Detection via Triangular Random Walks</b>	Boldi	LAWEB	LA-WEB 2012: 48-56 (2012)			2012	LA-WEB 2012: 48-56 (2012)	ibid
43	<b>Robustness of social and web graphs to node removal"</b>	Vigna	Soc. Network Anal. Mining	1-14 (2012)	Springer	Germany	2012	1-14 (2012)	ibid
44	<b>Injecting Uncertainty in Graphs for Identity Obfuscation</b>	Vigna	PVLDB	PVLDB 5(11): 1376-1387 (2012)			2012	PVLDB 5(11): 1376-1387 (2012)	ibid
45	<b>Axioms for centrality</b>	Vigna	Internet Math	v.10(3-4), p.22 (2014)	Taylor-Francis	EU	2014	v.10(3-4), p.22 (2014)	ibid
46	<b>Cache-oblivious peeling of random hypergraphs</b>	Vigna	IEEE DCC 2014	(DCC 2014), IEEE pp.352-361. (2014)	IEEE	USA	2014	(DCC 2014), IEEE pp.352-361. (2014)	ibid
47	<b>Entity-linking via graph-distance minimization</b>	Proc. 3rd GRAPH Insp. Trav.		pp.30-43 (2014)		France	2014	pp.30-43 (2014)	ibid

		Engin. GRAPH ITE 2014								
--	--	--------------------------------	--	--	--	--	--	--	--	--

All 73 papers and preprints with their references are available at the project website <http://www.quantware.ups-tlse.fr/FETNADINE/pub.html>

TEMPLATE A2: LIST OF DISSEMINATION ACTIVITIES								
NO.	Type of activities <sup>3</sup>	Main leader	Title	Date/Period	Place	Type of audience <sup>4</sup>	Size of audience	Countries addressed
1	NADINE Workshop	Shepel yansky	<i>Spectral Properties of Complex Networks</i>	23 - 27 July 2012	ECT Trento Italy	international	40 scientists	World
2	NADINE workshop	Vigna	Directed Networks Days 2013	13 - 14 June 2013	Di Univ di Milano, Italy	international	20 scientists	EU
3	NADINE workshop	Benczur	Directed Networks Days 2014	8 - 11 May 2014	MTA SZTAKI Budapest	international	20 scientists	EU

<sup>3</sup> A drop down list allows choosing the dissemination activity: publications, conferences, workshops, web, press releases, flyers, articles published in the popular press, videos, media briefings, presentations, exhibitions, thesis, interviews, films, TV clips, posters, Other.

<sup>4</sup> A drop down list allows choosing the type of public: Scientific Community (higher education, Research), Industry, Civil Society, Policy makers, Medias, Other ('multiple choices' is possible).

4	Luchon summer school	Shepleyan sky	Network analysis and applications	21 June - 5 July 2014	Luchon France	international	20 scientists and 30 students	world

*The list of 67 conferences is available at <http://www.quantware.ups-tlse.fr/FETNADINE/confs.html>*



**Section B (Confidential<sup>5</sup> or public: confidential information to be marked clearly)**

**Part B1 => no patents**

The applications for patents, trademarks, registered designs, etc. shall be listed according to the template B1 provided hereafter.

The list should, specify at least one unique identifier e.g. European Patent application reference. For patent applications, only if applicable, contributions to standards should be specified. This table is cumulative, which means that it should always show all applications from the beginning until after the end of the project.

<b>TEMPLATE B1: LIST OF APPLICATIONS FOR PATENTS, TRADEMARKS, REGISTERED DESIGNS, ETC.</b>					
Type of IP Rights <sup>6</sup> :	Confidential Click on YES/NO	Foreseen embargo date dd/mm/yyyy	Application reference(s) (e.g. EP123456)	Subject or title of application	Applicant (s) (as on the application)

<sup>5</sup> Note to be confused with the "EU CONFIDENTIAL" classification for some security research projects.

<sup>6</sup> A drop down list allows choosing the type of IP rights: Patents, Trademarks, Registered designs, Utility models, Others.

## Part B2

Please complete the table hereafter: **see section “The potential impact”**

Type of Exploitable Foreground <sup>7</sup>	Description of exploitable foreground	Confidential Click on YES/NO	Foreseen embargo date dd/mm/yyyy	Exploitable product(s) or measure(s)	Sector(s) of application <sup>8</sup>	Timetable, commercial or any other use	Patents or other IPR exploitation (licences)	Owner & Other Beneficiary(s) involved

In addition to the table, please provide a text to explain the exploitable foreground, in particular:

- Its purpose
- How the foreground might be exploited, when and by whom
- IPR exploitable measures taken or intended
- Further research necessary, if any
- Potential/expected impact (quantify where possible)

---

<sup>7</sup><sup>19</sup> A drop down list allows choosing the type of foreground: General advancement of knowledge, Commercial exploitation of R&D results, Exploitation of R&D results via standards, exploitation of results through EU policies, exploitation of results through (social) innovation.

<sup>8</sup> A drop down list allows choosing the type sector (NACE nomenclature) : [http://ec.europa.eu/competition/mergers/cases/index/nace\\_all.html](http://ec.europa.eu/competition/mergers/cases/index/nace_all.html)

### 4.3 Report on societal implications

Replies to the following questions will assist the Commission to obtain statistics and indicators on societal and socio-economic issues addressed by projects. The questions are arranged in a number of key themes. As well as producing certain statistics, the replies will also help identify those projects that have shown a real engagement with wider societal issues, and thereby identify interesting approaches to these issues and best practices. The replies for individual projects will not be made public.

#### **A General Information** (completed automatically when *Grant Agreement number* is entered).

Grant Agreement Number:

Title of Project: NADINE

Name and Title of Coordinator:

#### **B Ethics**

<p><b>1. Did your project undergo an Ethics Review (and/or Screening)?</b></p> <ul style="list-style-type: none"> <li>If Yes: have you described the progress of compliance with the relevant Ethics Review/Screening Requirements in the frame of the periodic/final project reports?</li> </ul> <p>Special Reminder: the progress of compliance with the Ethics Review/Screening Requirements should be described in the Period/Final Project Reports under the Section 3.2.2 'Work Progress and Achievements'</p>	<b>No</b>
<p><b>2. Please indicate whether your project involved any of the following issues (tick box) :</b></p>	<b>no</b>
<b>RESEARCH ON HUMANS</b>	
<ul style="list-style-type: none"> <li>Did the project involve children?</li> </ul>	no
<ul style="list-style-type: none"> <li>Did the project involve patients?</li> </ul>	no
<ul style="list-style-type: none"> <li>Did the project involve persons not able to give consent?</li> </ul>	no
<ul style="list-style-type: none"> <li>Did the project involve adult healthy volunteers?</li> </ul>	no
<ul style="list-style-type: none"> <li>Did the project involve Human genetic material?</li> </ul>	no
<ul style="list-style-type: none"> <li>Did the project involve Human biological samples?</li> </ul>	no
<ul style="list-style-type: none"> <li>Did the project involve Human data collection?</li> </ul>	no
<b>RESEARCH ON HUMAN EMBRYO/FOETUS</b>	
<ul style="list-style-type: none"> <li>Did the project involve Human Embryos?</li> </ul>	no
<ul style="list-style-type: none"> <li>Did the project involve Human Foetal Tissue / Cells?</li> </ul>	no
<ul style="list-style-type: none"> <li>Did the project involve Human Embryonic Stem Cells (hESCs)?</li> </ul>	no

• Did the project on human Embryonic Stem Cells involve cells in culture?	no
• Did the project on human Embryonic Stem Cells involve the derivation of cells from Embryos?	no
<b>PRIVACY</b>	
• Did the project involve processing of genetic information or personal data (eg. health, sexual lifestyle, ethnicity, political opinion, religious or philosophical conviction)?	no
• Did the project involve tracking the location or observation of people?	no
<b>RESEARCH ON ANIMALS</b>	
• Did the project involve research on animals?	no
• Were those animals transgenic small laboratory animals?	no
• Were those animals transgenic farm animals?	no
• Were those animals cloned farm animals?	no
• Were those animals non-human primates?	no
<b>RESEARCH INVOLVING DEVELOPING COUNTRIES</b>	
• Did the project involve the use of local resources (genetic, animal, plant etc)?	No
• Was the project of benefit to local community (capacity building, access to healthcare, education etc)?	No
<b>DUAL USE</b>	
• Research having direct military use	No
• Research having the potential for terrorist abuse	No

## C Workforce Statistics

3. Workforce statistics for the project: Please indicate in the table below the number of people who worked on the project (on a headcount basis).

Type of Position	Number of Women	Number of Men
Scientific Coordinator		1
Work package leaders	1	2
Experienced researchers (i.e. PhD holders)		5
PhD Students	1	
Other		

4. How many additional researchers (in companies and universities) were recruited specifically for this project?	3
Of which, indicate the number of men:	

<b>D Gender Aspects</b>		
<b>5. Did you carry out specific Gender Equality Actions under the project?</b>	<input type="radio"/> * <input type="radio"/>	<input type="checkbox"/> Yes <input type="checkbox"/> No
<b>6. Which of the following actions did you carry out and how effective were they?</b>		
	<b>Not at all effective</b>	<b>Very effective</b>
Design and implement an equal opportunity policy		<input type="radio"/>
Set targets to achieve a gender balance in the workforce		<input type="radio"/>
Organise conferences and workshops on gender		<input type="radio"/>
Actions to improve work-life balance		<input type="radio"/>
Other: <input type="text" value="No specific actions"/>		
<b>7. Was there a gender dimension associated with the research content – i.e. wherever people were the focus of the research as, for example, consumers, users, patients or in trials, was the issue of gender considered and addressed?</b>		
Yes- please specify	<input type="text"/>	
No X		
<b>E Synergies with Science Education</b>		
<b>8. Did your project involve working with students and/or school pupils (e.g. open days, participation in science festivals and events, prizes/competitions or joint projects)?</b>		
Yes- please specify	<input type="text"/>	
No X		
<b>9. Did the project generate any science education material (e.g. kits, websites, explanatory booklets, DVDs)?</b>		
Yes- please specify: video lecture at ecole de Luchon see project website	<input type="text"/>	
No		
<b>F Interdisciplinarity</b>		
<b>10. Which disciplines (see list below) are involved in your project?</b>		
Main discipline <sup>2</sup> : 1.1		
Associated disciplin: 1.2; 5.2	<input type="radio"/>	Associated discipline:
<b>G Engaging with Civil society and policy makers</b>		
<b>11a Did your project engage with societal actors beyond the research community? (if 'No', go to Question 14)</b>	<input type="radio"/> <input type="radio"/>	<input type="checkbox"/> Yes <input type="checkbox"/> No
<b>11b If yes, did you engage with citizens (citizens' panels / juries) or organised civil society (NGOs, patients' groups etc.)?</b>		

<p>No X</p> <p>Yes- in determining what research should be performed</p> <p>Yes - in implementing the research</p> <p>Yes, in communicating /disseminating / using the results of the project</p>				
<p><b>11c In doing so, did your project involve actors whose role is mainly to organise the dialogue with citizens and organised civil society (e.g. professional mediator; communication company, science museums)?</b></p>			<p><input type="radio"/></p> <p><input checked="" type="radio"/></p>	<p>Yes</p> <p>No</p>
<p><b>12. Did you engage with government / public bodies or policy makers (including international organisations)</b></p>				
<p>No</p> <p>Yes- in framing the research agenda</p> <p>Yes - in implementing the research agenda</p> <p>Yes, in communication: UN COMTRADE, WTO Geneve</p>				
<p><b>13a Will the project generate outputs (expertise or scientific advice) which could be used by policy makers?</b></p> <p>Yes – as a <b>primary</b> objective (please indicate areas below- multiple answers possible)</p> <p>Yes – as a <b>secondary</b> objective recommendations for international world trade UN COMTRADE, OECD_WTO</p> <p>No</p>				
<p><b>13b If Yes, in which fields?</b></p>				
<p><a href="#">Agriculture</a></p> <p><a href="#">Audiovisual and Media</a></p> <p><a href="#">Budget</a></p> <p><a href="#">Competition</a></p> <p><a href="#">Consumers</a></p> <p><a href="#">Culture</a></p> <p><a href="#">Customs</a></p> <p><a href="#">Development Economic and Monetary Affairs</a></p> <p><a href="#">Education, Training, Youth</a></p> <p><a href="#">Employment and Social Affairs</a></p>		<p><a href="#">Energy</a></p> <p><a href="#">Enlargement</a></p> <p><a href="#">Enterprise</a></p> <p><a href="#">Environment</a></p> <p><a href="#">External Relations</a></p> <p><a href="#">External Trade</a></p> <p><a href="#">Fisheries and Maritime Affairs</a></p> <p><a href="#">Food Safety</a></p> <p><a href="#">Foreign and Security Policy</a></p> <p><a href="#">Fraud</a></p> <p><a href="#">Humanitarian aid</a></p>	<p><a href="#">Human rights</a></p> <p><a href="#">Information Society</a></p> <p><a href="#">Institutional affairs</a></p> <p><a href="#">Internal Market</a></p> <p><a href="#">Justice, freedom and security</a></p> <p><a href="#">Public Health</a></p> <p><a href="#">Regional Policy</a></p> <p><a href="#">Research and Innovation</a></p> <p>Space</p> <p><a href="#">Taxation</a></p> <p><a href="#">Transport</a></p>	

<b>13c If Yes, at which level?</b>		
Local / regional levels		
National level		
European level		
International level		
<b>H Use and dissemination</b>		
<b>14. How many Articles were published/accepted for publication in peer-reviewed journals?</b>	<b>73</b>	
<b>To how many of these is open access<sup>3</sup> provided?</b>	<b>73 via arxiv and NADINE website</b>	
<b>How many of these are published in open access journals?</b>	<b>4</b>	
<b>How many of these are published in open repositories?</b>		
<b>To how many of these is open access not provided?</b>		
<b>Please check all applicable reasons for not providing open access:</b>		
<input type="checkbox"/> publisher's licensing agreement would not permit publishing in a repository <input type="checkbox"/> no suitable repository available <input type="checkbox"/> no suitable open access journal available <input type="checkbox"/> no funds available to publish in an open access journal <input type="checkbox"/> lack of time and resources <input type="checkbox"/> lack of information on open access <input type="checkbox"/> other <sup>4</sup> : .....		
<b>15. How many new patent applications ('priority filings') have been made?</b> <i>("Technologically unique": multiple applications for the same invention in different jurisdictions should be counted as just one application of grant).</i>		
<b>16. Indicate how many of the following Intellectual Property Rights were applied for (give number in each box).</b>	Trademark	
	Registered design	
	Other	

<sup>3</sup> Open Access is defined as free of charge access for anyone via Internet.

<sup>4</sup> For instance: classification for security project.

17. How many spin-off companies were created / are planned as a direct result of the project?		0
<i>Indicate the approximate number of additional jobs in these companies:</i>		
18. Please indicate whether your project has a potential impact on employment, in comparison with the situation before your project:		
Increase in employment, or	<input type="checkbox"/>	In small & medium-sized enterprises
Safeguard employment, or	<input type="checkbox"/>	In large companies
Decrease in employment,	<input type="checkbox"/>	None of the above / not relevant to the project
Difficult to estimate / not possible to quantify	<input checked="" type="checkbox"/>	
19. For your project partnership please estimate the employment effect resulting directly from your participation in Full Time Equivalent (FTE = one person working fulltime for a year) jobs:		Indicate figure:
Difficult to estimate / not possible to quantify		<input checked="" type="checkbox"/>
<b>I Media and Communication to the general public</b>		
20. As part of the project, were any of the beneficiaries professionals in communication or media relations?		
Yes <input checked="" type="checkbox"/> (see text above)		No <input type="checkbox"/>
21. As part of the project, have any beneficiaries received professional media / communication training / advice to improve communication with the general public?		
Yes <input type="checkbox"/>		No <input checked="" type="checkbox"/>
22. Which of the following have been used to communicate information about your project to the general public, or have resulted from your project?		
Press Release	<input checked="" type="checkbox"/>	Coverage in specialist press
Media briefing	<input type="checkbox"/>	Coverage in general (non-specialist) press
TV coverage / report	<input type="checkbox"/>	Coverage in national press <input checked="" type="checkbox"/>
Radio coverage / report	<input type="checkbox"/>	Coverage in international press <input checked="" type="checkbox"/>
Brochures /posters / flyers	<input type="checkbox"/>	Website for the general public / internet <input checked="" type="checkbox"/>
DVD /Film /Multimedia	<input type="checkbox"/>	Event targeting general public (festival, conference, exhibition, science café) <input checked="" type="checkbox"/>
23. In which languages are the information products for the general public produced?		
Language of the coordinator	<input type="checkbox"/>	English
Other language(s)		

**Question F-10:** Classification of Scientific Disciplines according to the Frascati Manual 2002 (Proposed Standard Practice for Surveys on Research and Experimental Development, OECD 2002):

**FIELDS OF SCIENCE AND TECHNOLOGY**

1.          NATURAL SCIENCES



- 1.1 Mathematics and computer sciences [mathematics and other allied fields: computer sciences and other allied subjects (software development only; hardware development should be classified in the engineering fields)] X
  - 1.2 Physical sciences (astronomy and space sciences, physics and other allied subjects)
  - 1.3 Chemical sciences (chemistry, other allied subjects)
  - 1.4 Earth and related environmental sciences (geology, geophysics, mineralogy, physical geography and other geosciences, meteorology and other atmospheric sciences including climatic research, oceanography, vulcanology, palaeoecology, other allied sciences)
  - 1.5 Biological sciences (biology, botany, bacteriology, microbiology, zoology, entomology, genetics, biochemistry, biophysics, other allied sciences, excluding clinical and veterinary sciences)
2. ENGINEERING AND TECHNOLOGY
- 2.1 Civil engineering (architecture engineering, building science and engineering, construction engineering, municipal and structural engineering and other allied subjects)
  - 2.2 Electrical engineering, electronics [electrical engineering, electronics, communication engineering and systems, computer engineering (hardware only) and other allied subjects]
  - 2.3. Other engineering sciences (such as chemical, aeronautical and space, mechanical, metallurgical and materials engineering, and their specialised subdivisions; forest products; applied sciences such as geodesy, industrial chemistry, etc.; the science and technology of food production; specialised technologies of interdisciplinary fields, e.g. systems analysis, metallurgy, mining, textile technology and other applied subjects)
3. MEDICAL SCIENCES
- 3.1 Basic medicine (anatomy, cytology, physiology, genetics, pharmacy, pharmacology, toxicology, immunology and immunohaematology, clinical chemistry, clinical microbiology, pathology)
  - 3.2 Clinical medicine (anaesthesiology, paediatrics, obstetrics and gynaecology, internal medicine, surgery, dentistry, neurology, psychiatry, radiology, therapeutics, otorhinolaryngology, ophthalmology)
  - 3.3 Health sciences (public health services, social medicine, hygiene, nursing, epidemiology)
4. AGRICULTURAL SCIENCES
- 4.1 Agriculture, forestry, fisheries and allied sciences (agronomy, animal husbandry, fisheries, forestry, horticulture, other allied subjects)
  - 4.2 Veterinary medicine
5. SOCIAL SCIENCES
- 5.1 Psychology
  - 5.2 Economics
  - 5.3 Educational sciences (education and training and other allied subjects)
  - 5.4 Other social sciences [anthropology (social and cultural) and ethnology, demography, geography (human, economic and social), town and country planning, management, law, linguistics, political sciences, sociology, organisation and methods, miscellaneous social sciences and interdisciplinary , methodological and historical S1T activities relating to subjects in this group. Physical anthropology, physical geography and psychophysiology should normally be classified with the natural sciences].
6. HUMANITIES
- 6.1 History (history, prehistory and history, together with auxiliary historical disciplines such as archaeology, numismatics, palaeography, genealogy, etc.)
  - 6.2 Languages and literature (ancient and modern)
  - 6.3 Other humanities [philosophy (including the history of science and technology) arts, history of art, art criticism, painting, sculpture, musicology, dramatic art excluding artistic "research" of any kind, religion, theology, other fields and subjects pertaining to the humanities, methodological, historical and other S1T activities relating to the subjects in this group]

## 2. FINAL REPORT ON THE DISTRIBUTION OF THE EUROPEAN UNION FINANCIAL CONTRIBUTION

---

This report shall be submitted to the Commission within 30 days after receipt of the final payment of the European Union financial contribution.

### Report on the distribution of the European Union financial contribution between beneficiaries

Name of beneficiary	Final amount of EU contribution per beneficiary in Euros
1.	
2.	
n	
Total	